C T I

A Hybrid Architecture for Metacognitive Learning

Cognitive Technologies, Inc.

4200 Lorcom Lane

Arlington, VA 22207

Hybrid Architectures for Complex Learning Program

Contract No. N00014-95-C-0182 with the Office of Navy Research

COTR: Helen Gigley

13 August 1997

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
13 AUG 1997		**00-00-1997 to 00-00-1997**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
A Hybrid Architecture for Metacognitive Learning	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Cognitive Technologies Inc,4200 Lorcom Lane,Arlington,VA,22207**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Second Meeting, Accelerated Research Initiative on Hybrid Architectures as Models of Human Learning. San Diego, CA, 1997. U.S. Government or Federal Rights License

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**63**	

C T I

Research Issues

- **Investigate modeling of reflexive reasoning in the CIC task.**
- **Does metacognitive / reflective skill improve the efficiency and accuracy of reflexive learning?**
- **How is metacognitive skill itself learned?**

Initial machine learning study addresses the second issue.

C T I

Research Overview

1. **System Concept**
   - **Integrating *adaptable rapid parallel reflexive reasoning* with *adaptable executive processes for critiquing and correcting***
2. **SHRUTI Enhancements (Shastri)**
3. **Hybrid System Development**
   - **Knowledge Base Encoding**
   - **Metacognitive System**
   - **Adaptive Critic architecture**
4. **Machine Learning Study**
   - **Learning abductive reasoning by localized backpropagation**
   - **Effect of a metacognitive rule on domain learning**
5. **Human Learning Studies**
6. **Future Plans**

C T I

1. Hybrid Architecture for Metacognitive Learning

An Integration of Three Research Areas

- **Recognition / Metacognition model of critical thinking in decision making.**
- **SHRUTI, a connectionist model of rapid parallel reasoning with dynamic variable binding**
- **Adaptive Critics, a connectionist model of behavior learning.**

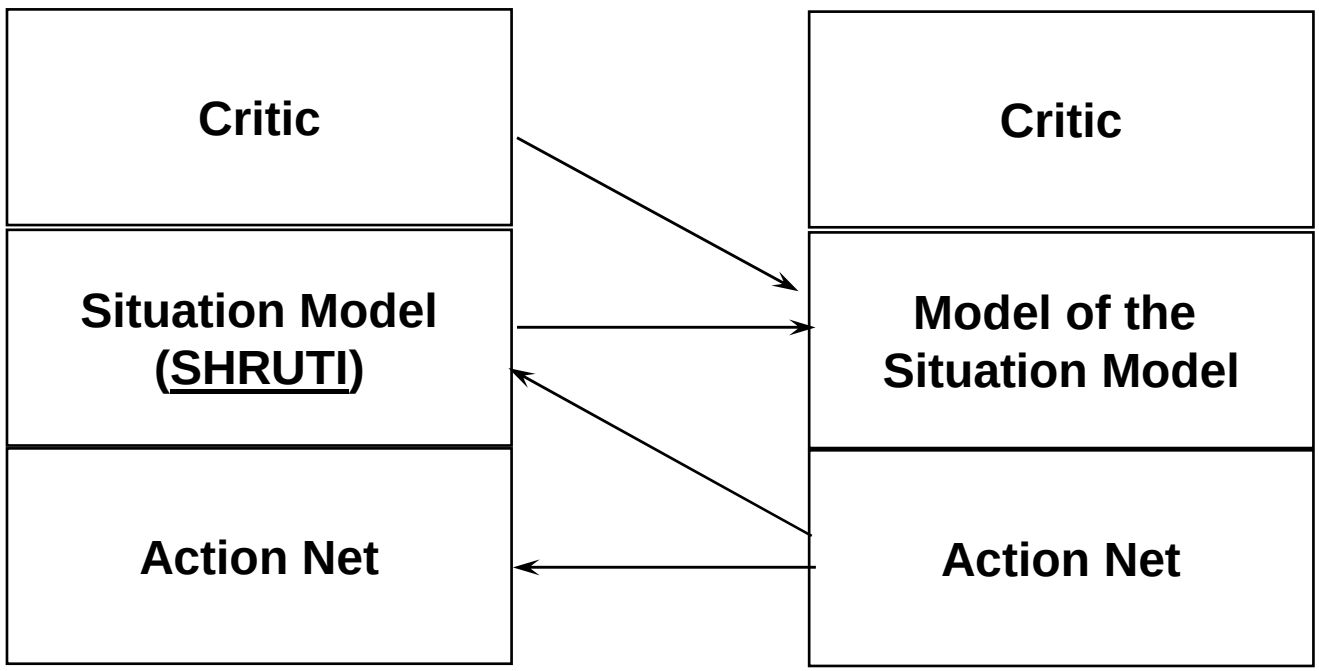
C T I

How Does the Architecture Relate These Three Components?

Recognition Adaptive Critic

Metacognition Adaptive Critic

Implements Recognition / Metacognition Model

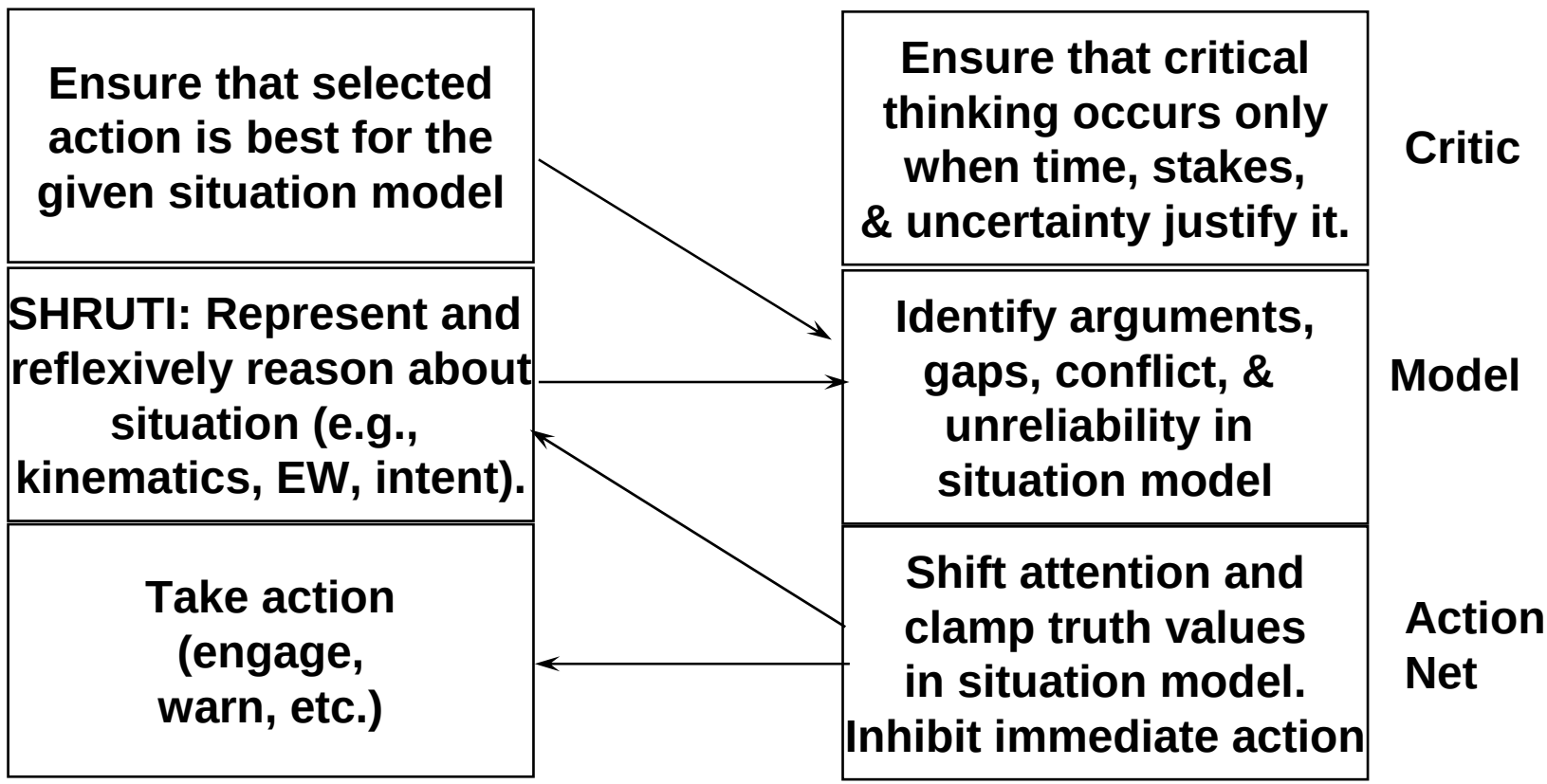


C T I

Functionality of Components

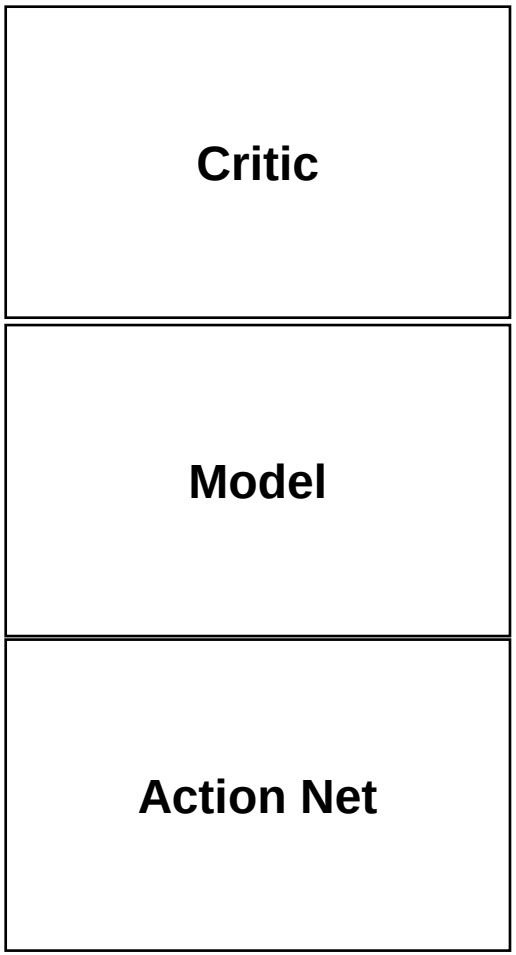
Recognition Adaptive Critic

Metacognitive Adaptive Critic (Implements R/M Model)



C T I

Hybrid System: Learning in Adaptive Critics



Critic learns to predict cumulative expected reinforcement associated with state or state and action

Model learns good predictions of valued events.

Action Net learns to reflexively take actions that maximize expected value of outcome.

C T I

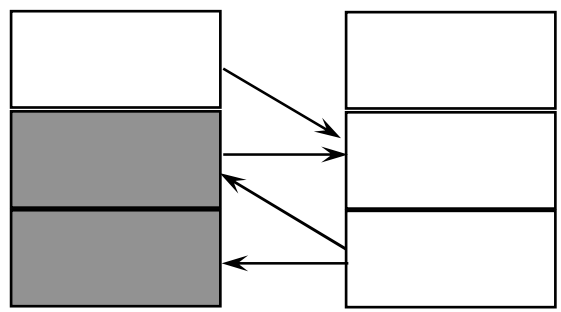
Summary of Hybrid Concept

- **Hybrid model captures both reflexive (recognitional) processes and metacognition (critical thinking)**
- **The combination of these is what enables decision makers both to exploit their experience and to go beyond it**
- **Metacognition leverages existing knowledge in novel and uncertain situations**
- **Metacognition increases efficiency of domain learning**

C T I

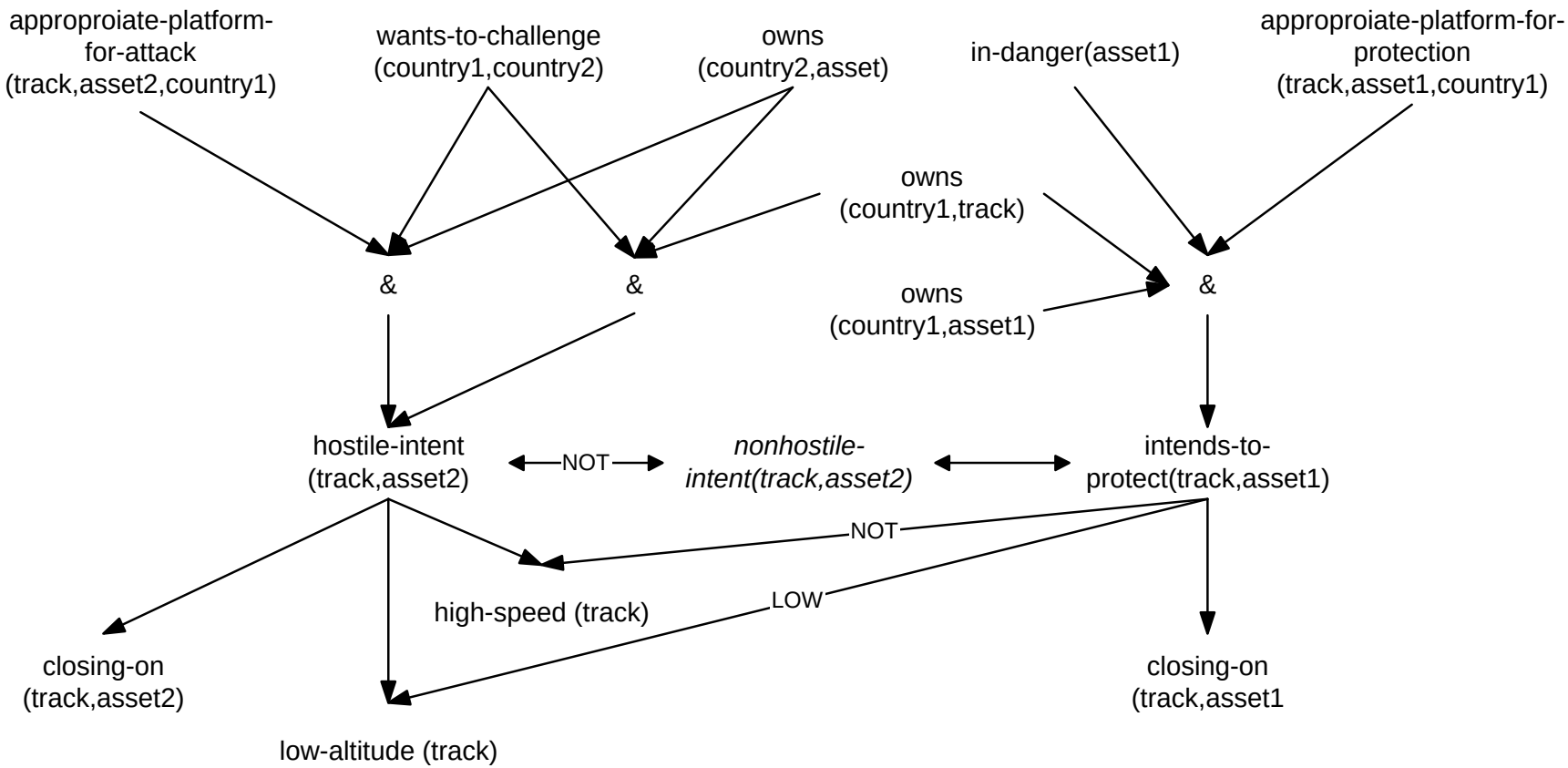
Knowledge Base Encoding In SHRUTI

- **Sources**
  - **Critical incident interviews with active-duty naval officers**
  - **Participants of experimental trials at Surface Warfare Officers School and Naval Post Graduate School**



C T I

Intent Story Structures



C T I

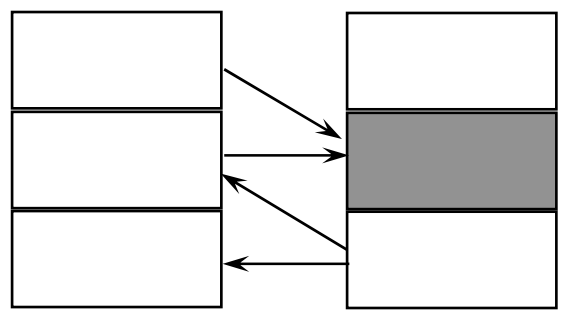
SHRUTI Enhancements

- **Negation**
- **Abductive reasoning**
- **Backpropagation**
- **Many others**

C T I

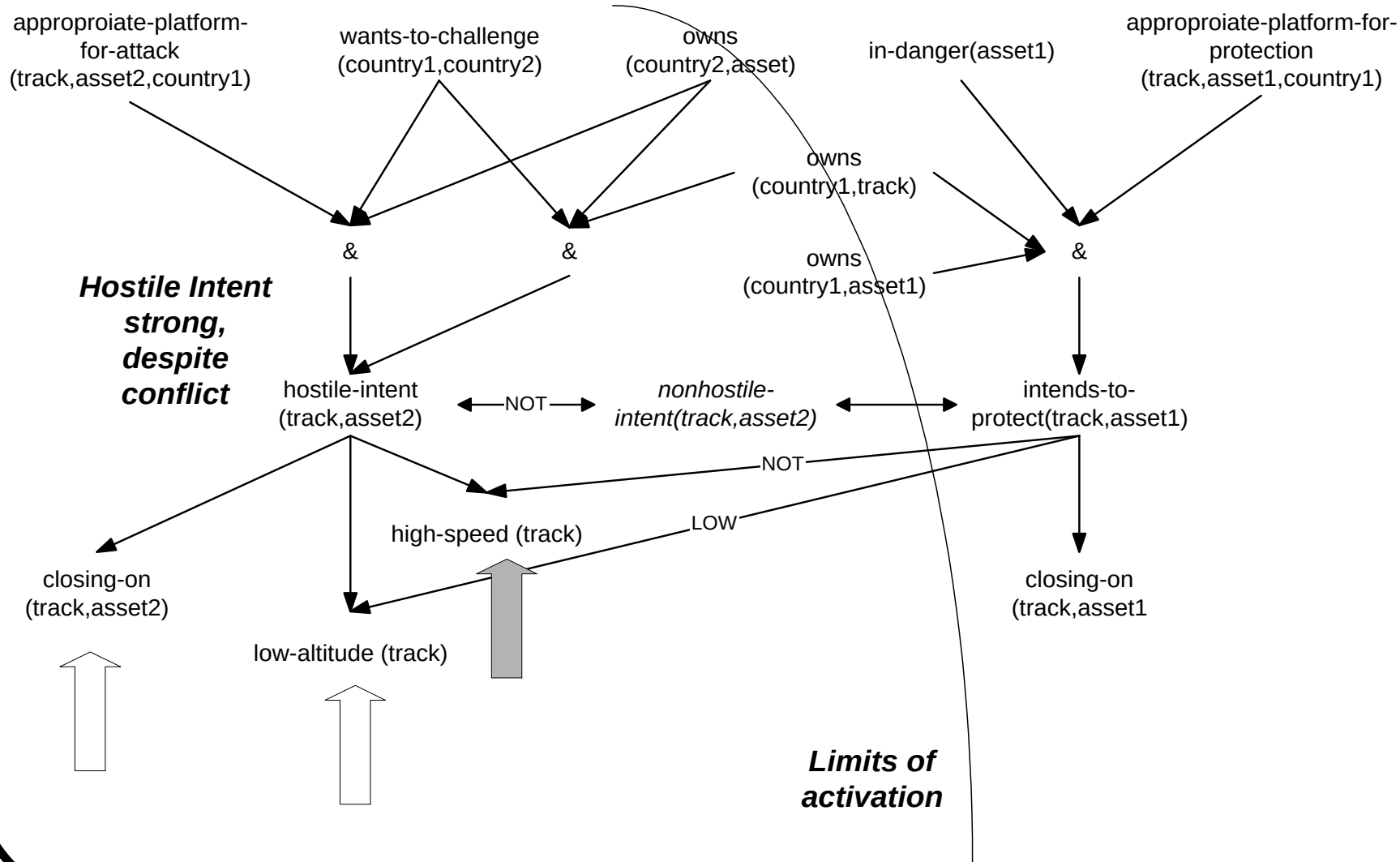
Metacognitive Model in Hybrid System

- **Bound on length of reasoning chain and limit on number of objects in SHRUTI means all relevant information in long term knowledge base may not be retrieved in first cycle of activation**
- **Metacognitive model provides mechanism for managing recognitional processing by shifting attention and clamping values**
- **Metacognitive critic balances advantages of thinking more versus risks of delaying action**



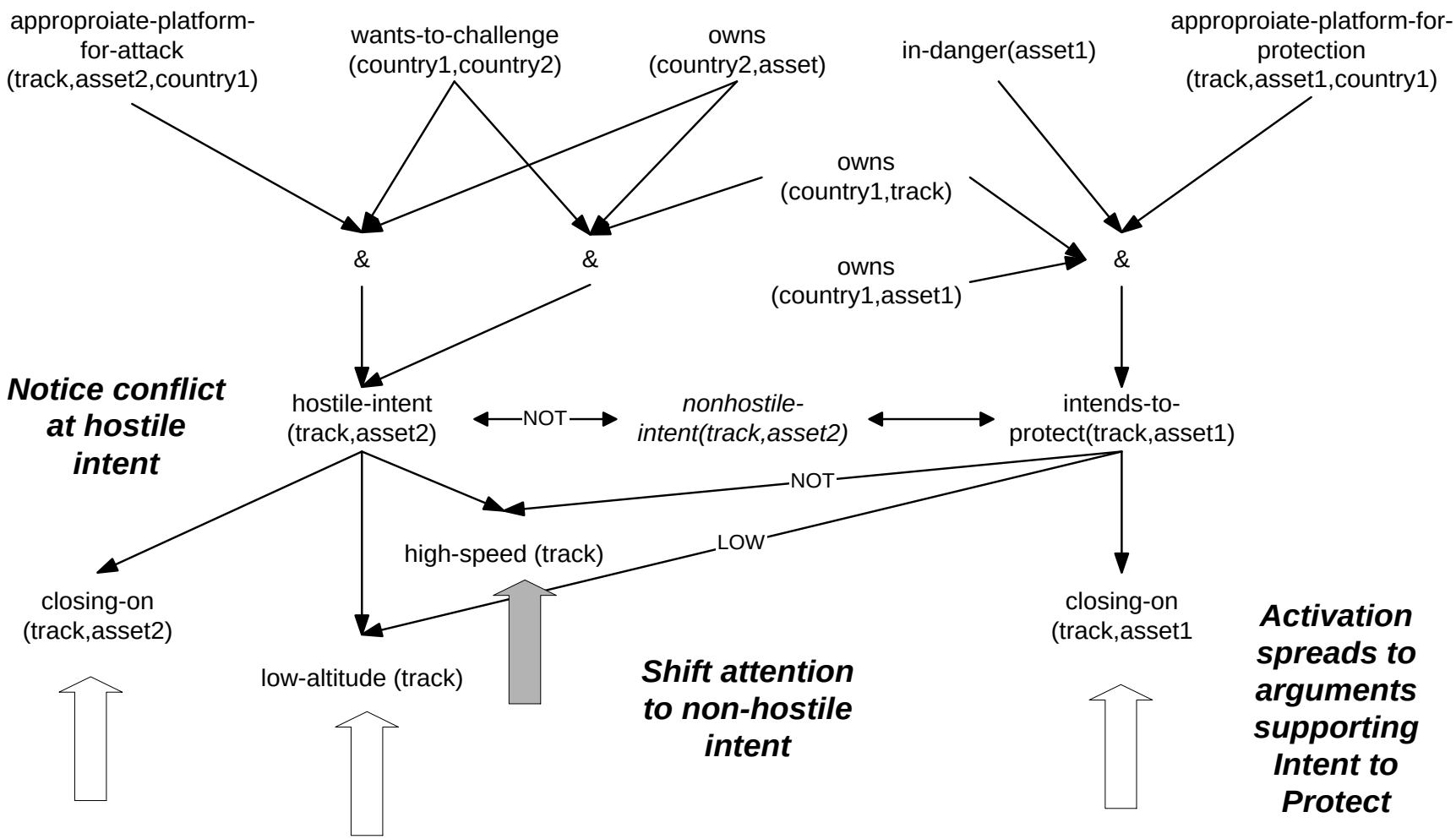
C T I

Limits on Reflexive Reasoning



C T I

Metacognitive Attention Shifting



C T I

Machine Demonstration

Reflexive performance

- **hostile intent + 675, - 169, conflict 114**
- **intent to protect + 0, - 394, conflict 0**
- **less experienced officer might engage**

Performance with metacognitive skill

- **hostile intent + 733, - 168, conflict 123**
- **intent to protect + 415, - 0, conflict 0**
- **more experienced officer might wait**

C T I

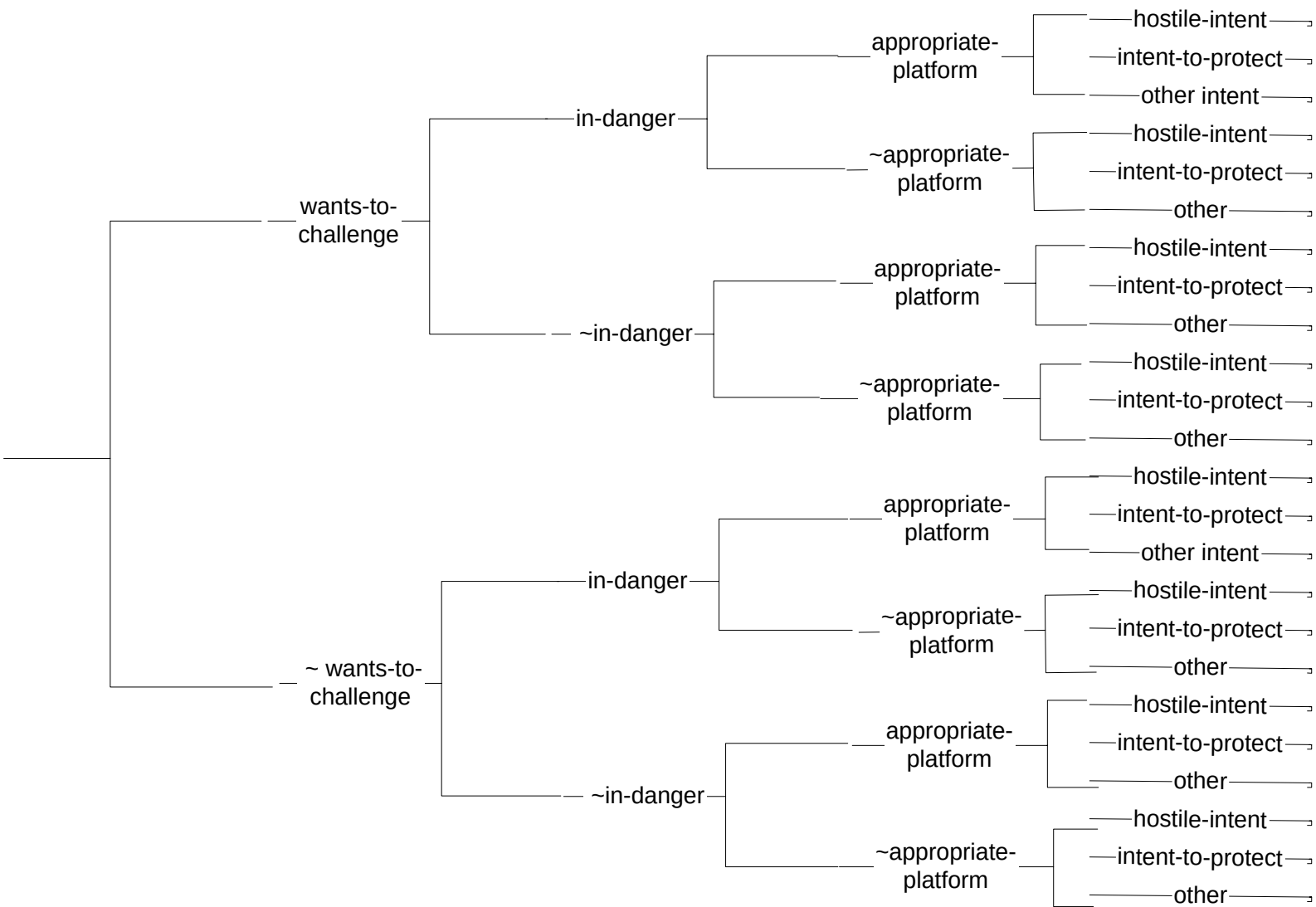
4. Machine Learning Study

- **Procedure**
  - **Generate scenarios according to “real-world” probabilities in event tree**
  - **Randomly divide scenarios into training and test sets**
  - **Use backpropagation to adapt weights in training scenarios**
  - **In reflexive condition, train each scenario for 50 cycles**
  - **In metacognitive condition, insert “hint” to query non-hostile intent midway through training**

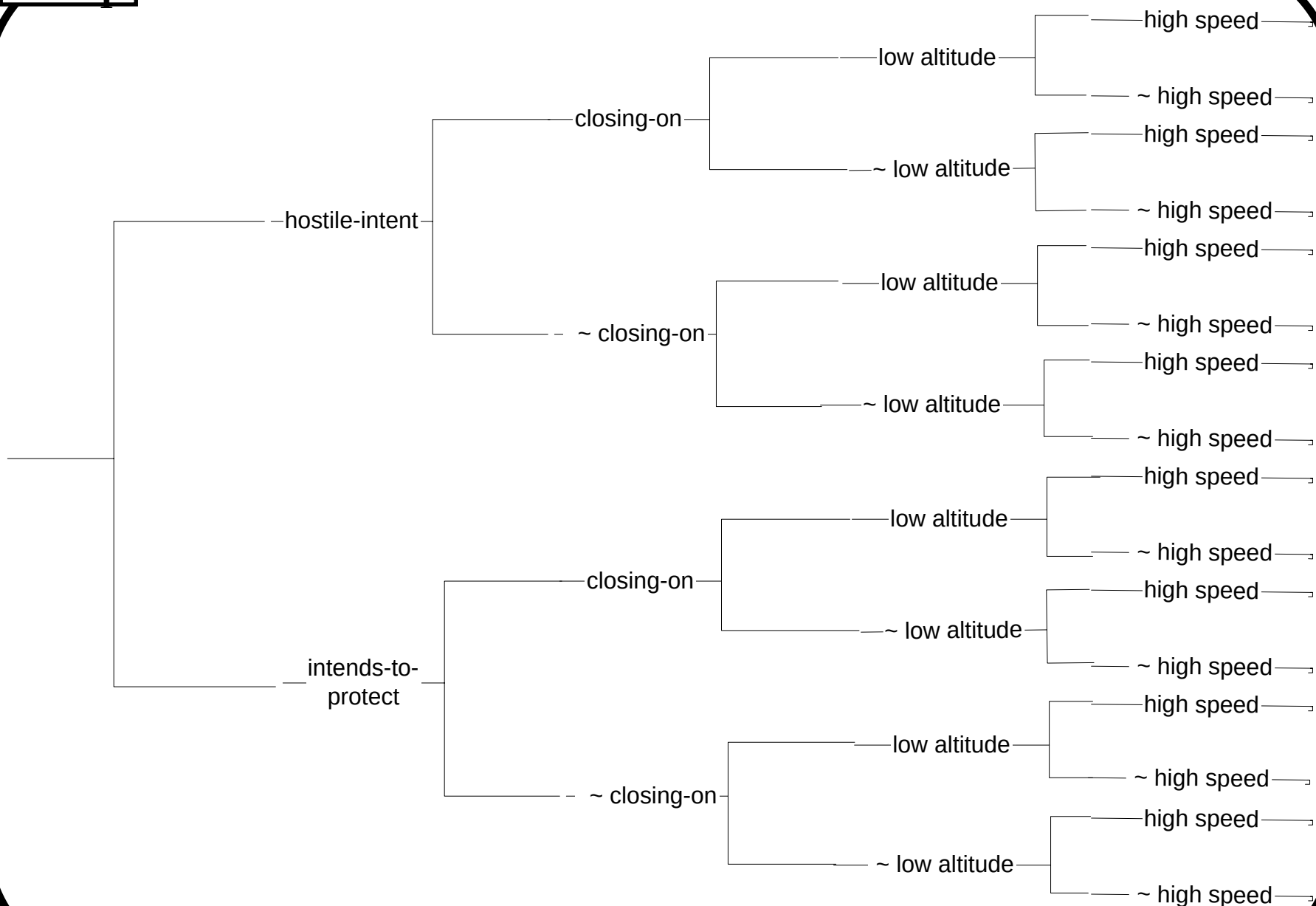
- **Dependent Variables**
  - **Changes in weights (I.e., knowledge base adaptation)**
  - **Probability-weighted sum of squared errors in test scenarios**
  - **Performance in standard scenario (Korea)**

C T I

Scenario Generation: Probabilistic Event Tree



C T I


hostile-intent
intends-to-protect
closing-on
~ closing-on
low altitude
~ low altitude
high speed
~ high speed

C T I

Demonstration: Effect of Metacognitive "Hint" During Training

Reflexive Training: Correction applied to hostile intent

- **Change in weight on rule for hostile intent from 900 to 750**

Training with metacognition: Correction as above, plus "hint" to think about non-hostile intents

- **Change in weight on rule for hostile intent from 900 to 750**
- **Increase in "prior belief" in likelihood of non-hostile intent from 50 to 64**
- **Decrease in "prior belief" in likelihood of hostile intent from 50 to 49**

C T I

Utility Function for Reflexive Critic

- **U(mil,intent,act) + U(attacked,mil) + U(ROEviol)**
- **Illustrative values**

U(mil,intent,act)

	Military	Civilian
Hostile intent	**Engage 20 Not eng 0**	**Engage 10 Not eng 0**
No hostile intent	**Engage -30 Not eng 0**	**Engage -100 Not eng 0**

U(attacked,mil)

Military	Civilian
Attacked -150 Not attacked 0	**Attacked -75 Not attacked 0**

C T I

Utility Function for Metacognitive Critic

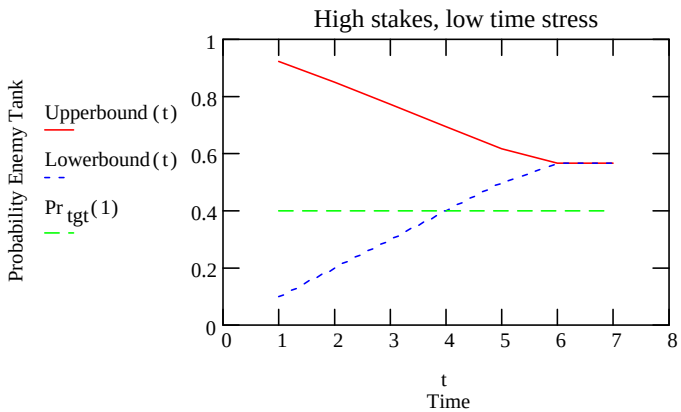
- **Change in payoffs due to change in actions due to change in beliefs**
- **Involves comparison of future utility expected given critical thinking versus utility of outcome that would occur if officer acted immediately on current best model**

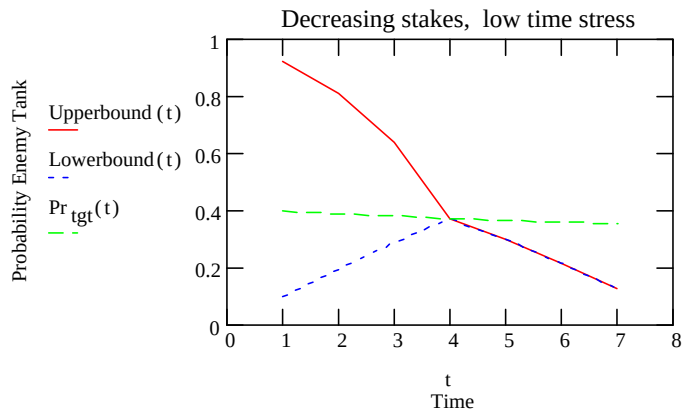
C T I

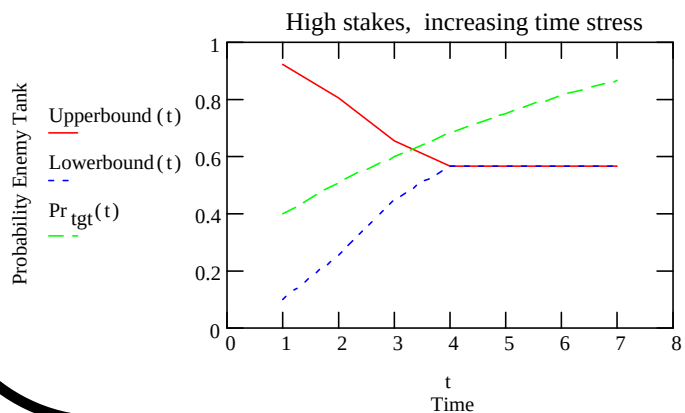
When Critical Thinking is Appropriate

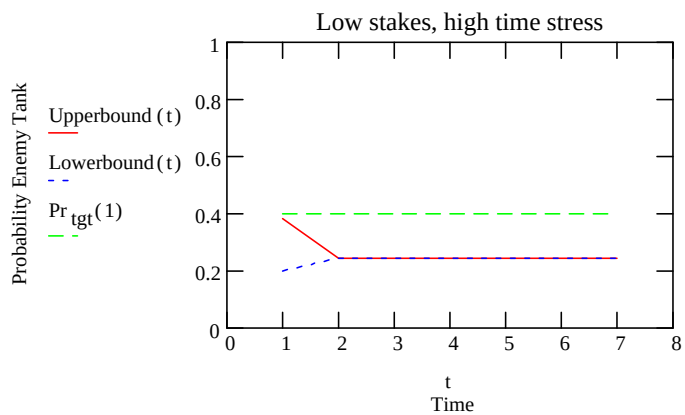
Decision to accept vs. question recognitional response

- Confidence in recognitional response
- Cost of delay
- Cost of error in acting on current best response









C T I

Handling Tracks: Complementary Schemas

ROE-EXECUTION SCHEMA	THREAT-ASSESSMENT SCHEMA
Start with tracks within 20 nm of ownship and move out in expanding circles to 30, 40, & 50 nm	Start with tracks within their weapons range of own ship, move out to tracks that can detect own ship
Query positive support for non-hostile(track,Ownship)	Query positive support for hostile-intent(track,Ownship)
Implement three levels of warning, locking on, and engagement	May lead to engagement outside narrow ROE criteria, or non-engagement when inside ROE criteria

C T I

Distinctive Aspects of Model Predictions

- **Reasoning depends critically on dynamic binding**
  - **Different countries own the two platforms in attack situation (attacking and attacked)**
  - **Same country owns two platforms in protection situation**
- **Includes background knowledge underlying "base rates"**
- **Applies to situations where all long-term knowledge is not immediately accessible due to cognitive limits**
- **Model predicts shifts of attention and resulting changes in belief**
- **Uncertainty is treated as a problem to be solved rather than averaged away**

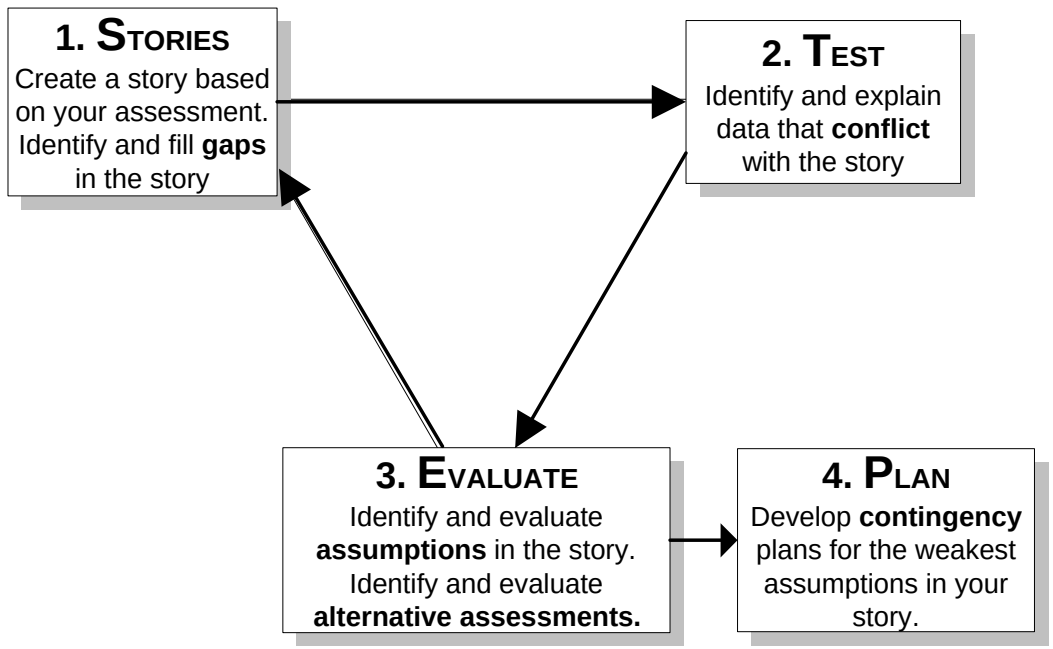
C T I

Effects of Metacognition on Long Term Memory Structure

- **General metacognitive skills**
  - **When and where to shift attention based on conflict, gaps, unreliability**
- **Domain-specific metacognitive skills**
  - **Which predicates to scrutinize when trouble arises**
- **Speeds compilation of domain knowledge**
  - **Extends reach of backpropagation learning**
  - **Constructs new links**

C T I

A Training Tool: The STEP Framework for Critical Thinking



C T I

Scenarios and Simulation

- **5 Scenarios developed by NCCOSC as part of the TADMUS program.**
- **Enhanced by CTI to emphasize features relevant to critical thinking**
  - **Information often matches no single pattern**
  - **Scenarios extended so that tracks come closer to own ship**

C T I

NPS Study

Participants

- **35 graduate students at the Naval Post Graduate School**
- **Median years of service = 9.5 years**

Experimental Design

Scenario Order x Test (Pre vs Post)

Procedure

- **Total of five 2-hour sessions, spread over two weeks**
- **Four hours of training**

C T I

Testing, Elaborating, & Evaluating Stories

- **Trained participants identified more conflicting evidence**
  - **NPS: 58% increase (p<.001)**
- **Trained officers identified more assumptions underlying conflicting evidence**
  - **NPS: 27% increase (p<.001**
- **Trained participants generated more alternative assessments**
  - **NPS: 41% increase (p<.001)**
- **Training significantly increased agreement with SME in two of the four test scenarios -- one at both NPS and SWOS (p = .034, p = .013; effect ns in the other scenarios)**

6. Future Plans

- **Reinforcement learning**
  - **Domain**
  - **Metacognitive**
- **Testing**
- **SHRUTI enhancements**

C T I

Machine Exploration of Theory

- **Determine relative importance of different metacognitive capabilities**
  - **Metacognitive actions: Shift focus versus clamp truth values**
  - **Metacognitive model elements: Identify arguments, gaps, conflict, &/or unreliability in knowledge base**
  - **Simulate and examine value of specific techniques, such as devil's advocate (crystal ball)**
- **Determine relative importance of general metacognitive rules versus domain-specific metacognitive rules**
  - » **Identifying arguments, gaps, conflict, unreliability**
  - » **Which arguments are examined first? Pattern and sequence of corrective actions. Standards for judging reliability. Methods for assessing risk and stakes**

C T I

Human & Machine Testing

Specific predictions of Hybrid Model

- **Benefits of metacognitive skill depend on limitations in length of inferential chains in reflexive reasoning**
- **Benefits of metacognitive skill increase with the number of objects to be reasoned about**
- **Order of generation of arguments and explanations corresponds to length of inferential chains**
- **Retrospective reinterpretation of previously considered evidence occurs as a result of new conflicting data**

C T I

Human & Machine Testing

- **Compare alternative training interventions**
  - **Practice**
    - » **Through practice, strategies should become more densely connected to KB and more fluidly executed (i.e., “compiled”)**
  - **Hints & feedback #1: Direct attention**
    - » **What is track #7031 doing?**
    - » **What do you think about the intermittent IFF response?**
    - » **Is that track a more lucrative target than own ship?**
  - **Hints & feedback #2: Clamp nodes true or false**
    - » **If the track is hostile, what would you expect to see?**
    - » **What else could that observation mean?**
    - » **Suppose you are wrong in your assessment. Explain how that could be.**

Appendix A.
SHRUTI

C T I

The Recognitional Model: SHRUTI

- **Rapid reflexive inference**
  - **Time independent of size of long term knowledge base**
  - **Space linear in size of long term knowledge base**
- **Neurally plausible dynamic variable binding**
  - **Goes beyond association of propositions or features (e,g., in schemas)**
  - **Supports complex relational reasoning among multiple objects, features, and relations**
- **Implies limitations on reflexive reasoning**
  - **Limits on number of objects & length of reasoning/retrieval chain, not on number of active facts or rules**
  - **Provides context in which metacognitive skills may improve both performance and learning**

C T I

Major SHRUTI Enhancements Completed

- **Negation**
  - **Support for negative as well as positive knowledge. Explains how agent can hold inconsistent beliefs, and become aware of contradiction only when beliefs are within a certain inferential distance of each other**
- **Abductive reasoning**
  - **Inference from occurrence of an effect to its likely cause; combination of evidence regarding a cause**
- **Derivative computation to support back propagation / reinforcement learning**

C T I

SHRUTI Enhancements Completed

- **Summary facts**
  - **Numerically represent strength of belief in a fact**
  - **Strength of belief affected by types of role fillers, e.g., a particular type of aircraft is more likely to attack cruiser than other types**
- **Reduced restrictions on rules**
  - **Regarding multiple occurrences of same variable in antecedent or consequent but not both**
- **Continuous attribute values**
  - **Representation of scalar quantities such as speed and altitude**
- **Action schemas**
  - **Procedural sequences**
- **Mutual exclusion in type hierarchy**
- **Ability to save snapshots of network weights and activity patterns**

C T I

Metacognitive Model

Representations of

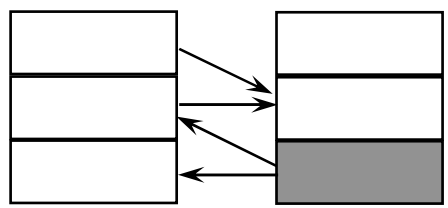
- **Attended arguments in reflexive model (conclusion-grounds pairs)**
- **Different types of uncertainty associated with arguments in reflexive model**
  - **Incompleteness The grounds of an argument activate neither + nor - collector of conclusion**
  - **Conflict: Grounds1 activates + and grounds2 activates - collector for conclusion — provided that query involves only instantiated or universally quantified variables**
  - **Unreliability: Grounds activate either + or - conclusion and not both, but activation is changeable due to as yet unconsidered arguments**

C T I

Metacognitive Action Net

Basic metacognitive actions

- **Shift attention / query predicates placed in intermediate memory by reflexive model**
  - **Training analog: “hints” to think about a possibility**
- **Clamp predicates true or false**
  - **Training analog: “hints” to imagine your conclusion is wrong, hints to look for alternative explanations or predictions**
- **Inhibit actions of reflexive system**
  - **Training analog:**



C T I

Appendix B.
Knowledge Base Encoding

Aspects of Knowledge Base

- **Active schemas for handling tracks**
  - **Scan tracks according to ROE and doctrine based responding**
  - **Assess tracks for threat status**
- **Intent Story Structures**
  - **Includes overall situation (goals, capabilities, opportunities) and preparatory actions (e.g., localization, track kinematics and EW)**
- **Structures for defeasible reasoning**
  - **Allows for rapid initial inference, then possible withdrawal of conclusion on further examination**

C T I

Reflexive Actions

- **External response**
  - **engage**
  - **jam**
  - **launch chaff**
  - **vector CAP**
  - **maneuver**
  - **illuminate**
  - **warn (levels 1,2,3)**
  - **query**
  - **offer help**
- **Internal response**
  - **activate AAW auto**
  - **prepare defenses**
  - **get targeting solution**
  - **request assets**
  - **designate hostile**
  - **set alert level**
  - **broadcast to battle force**
  - **notify superior(s)**
  - **monitor**
  - **review / study / recall**

C T I

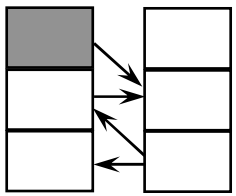
Types of Stories

- Attack
- Targeting
- Harass
- Provoke
- Testing response
- Reconnaissance
- Patrol
- Exercise
- Training
- Friendly
- Commair
- Search & rescue

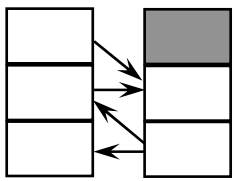
Appendix C
Reinforcement Learning in Adaptive Critics

C T I

Reinforcement Learning Accomplishments



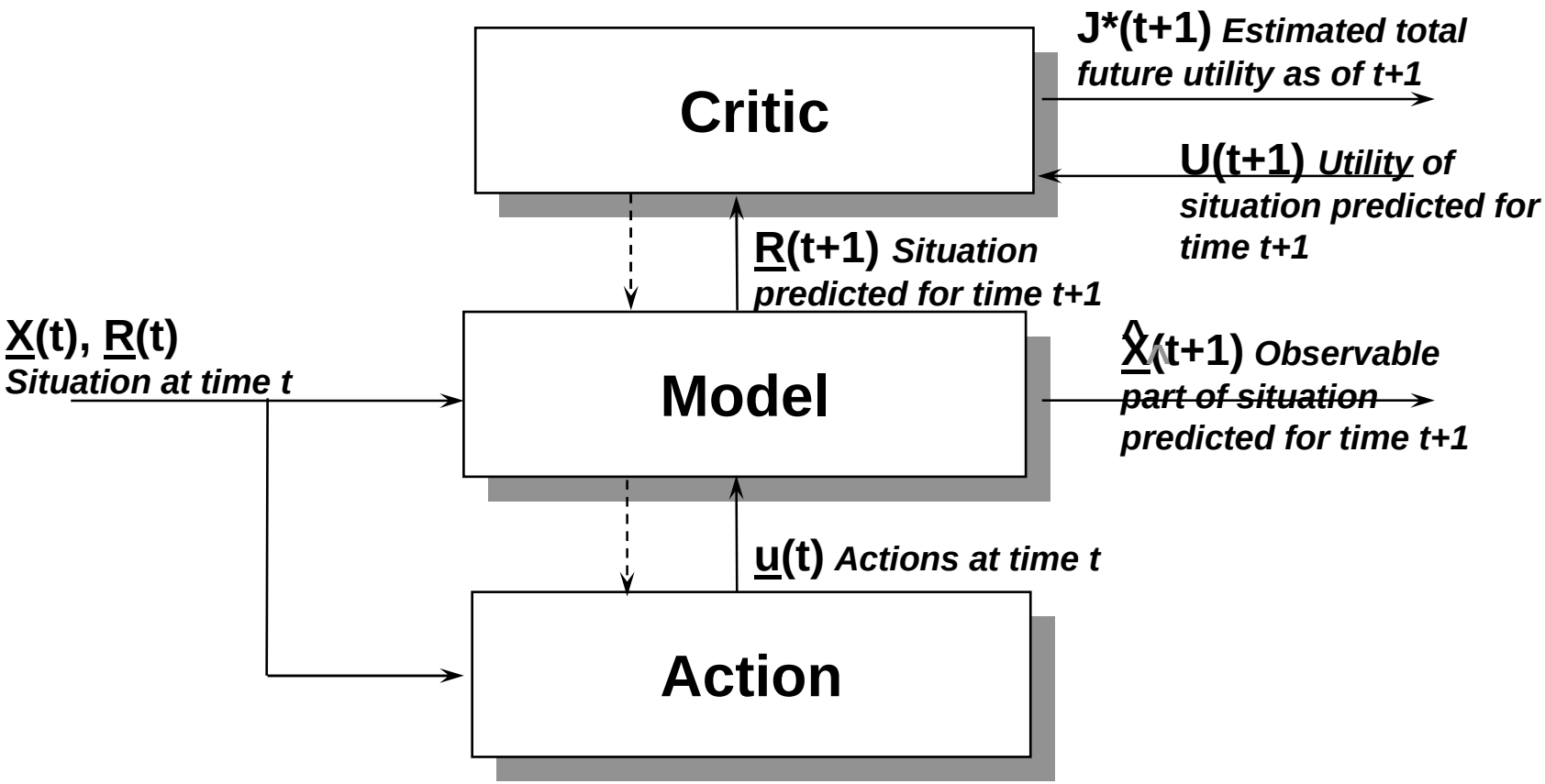
- **Specification of utility function for Recognition Critic**
  - **Represents expected value given the recognition model and selected action**
  - **Supports learning the best overt action given the scenario**



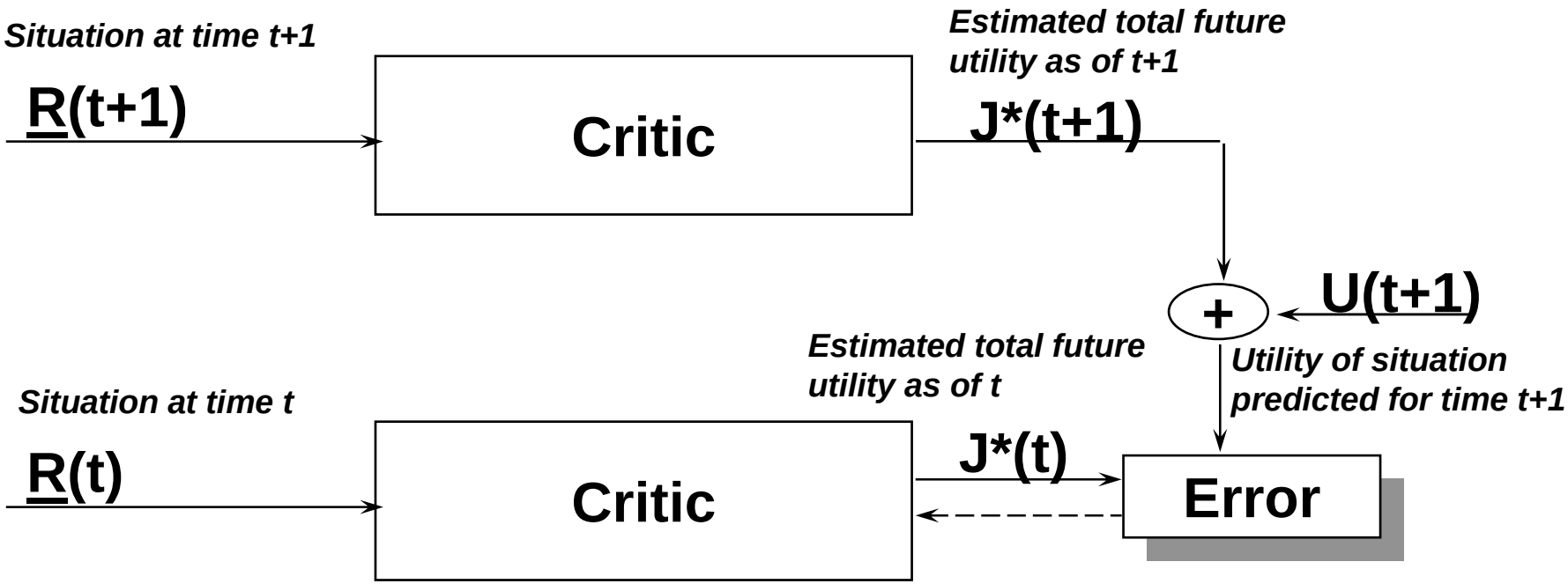
- **Specification of utility function for Metacognitive Critic**
  - **Represents expected value given the recognition model, the selected action, *and* the uncertainty in the recognition model**
  - ***S*upports learning the best metacognitive action given the recognition model**
- **Benchmark model of attention shifting**

C T I

Backpropagated Adaptive Critic (BAC)



Critic Learning

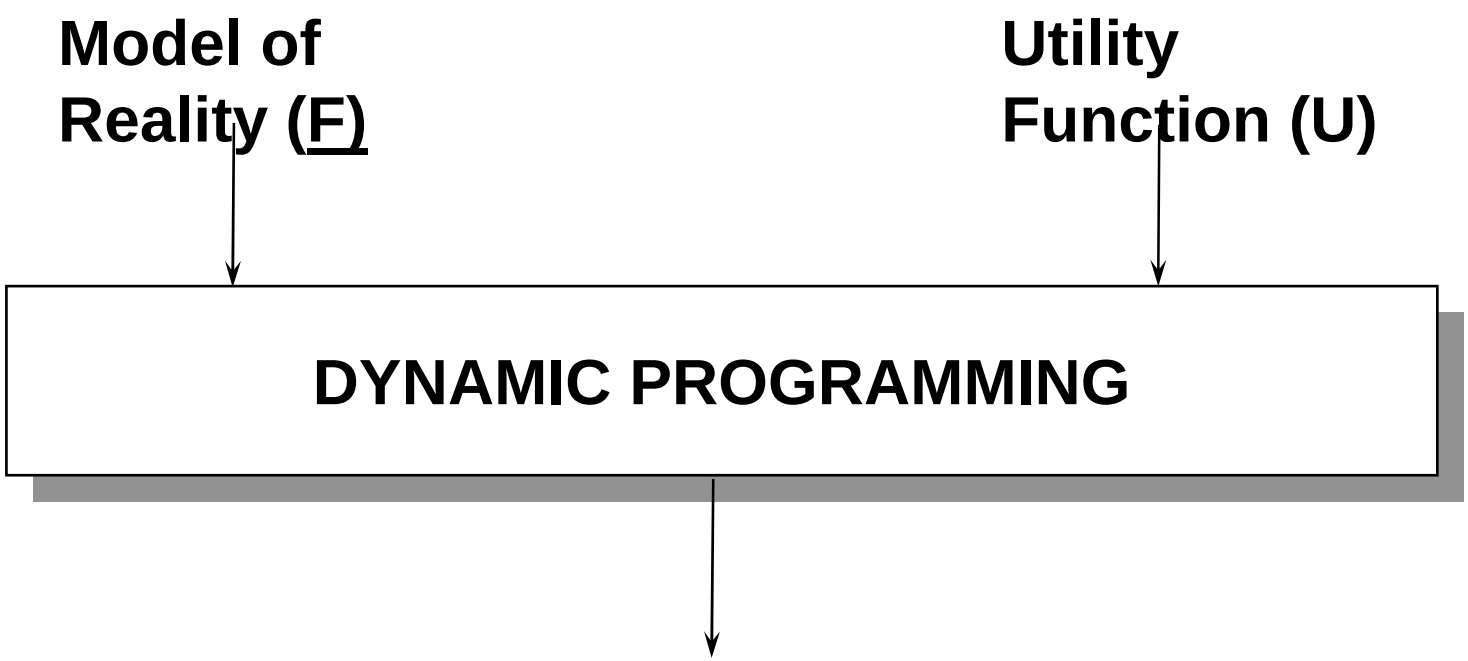


Adjust critic weights so that
J*(t+1) + U(t+1) = J*(t)

C T I

Dynamic Programming

- General, exact, and efficient
- Maximizes utility over time.
- Can control multiple variables.
- Works in noisy and nonlinear environments.



C T I

Benchmark Model of Attention Shifting

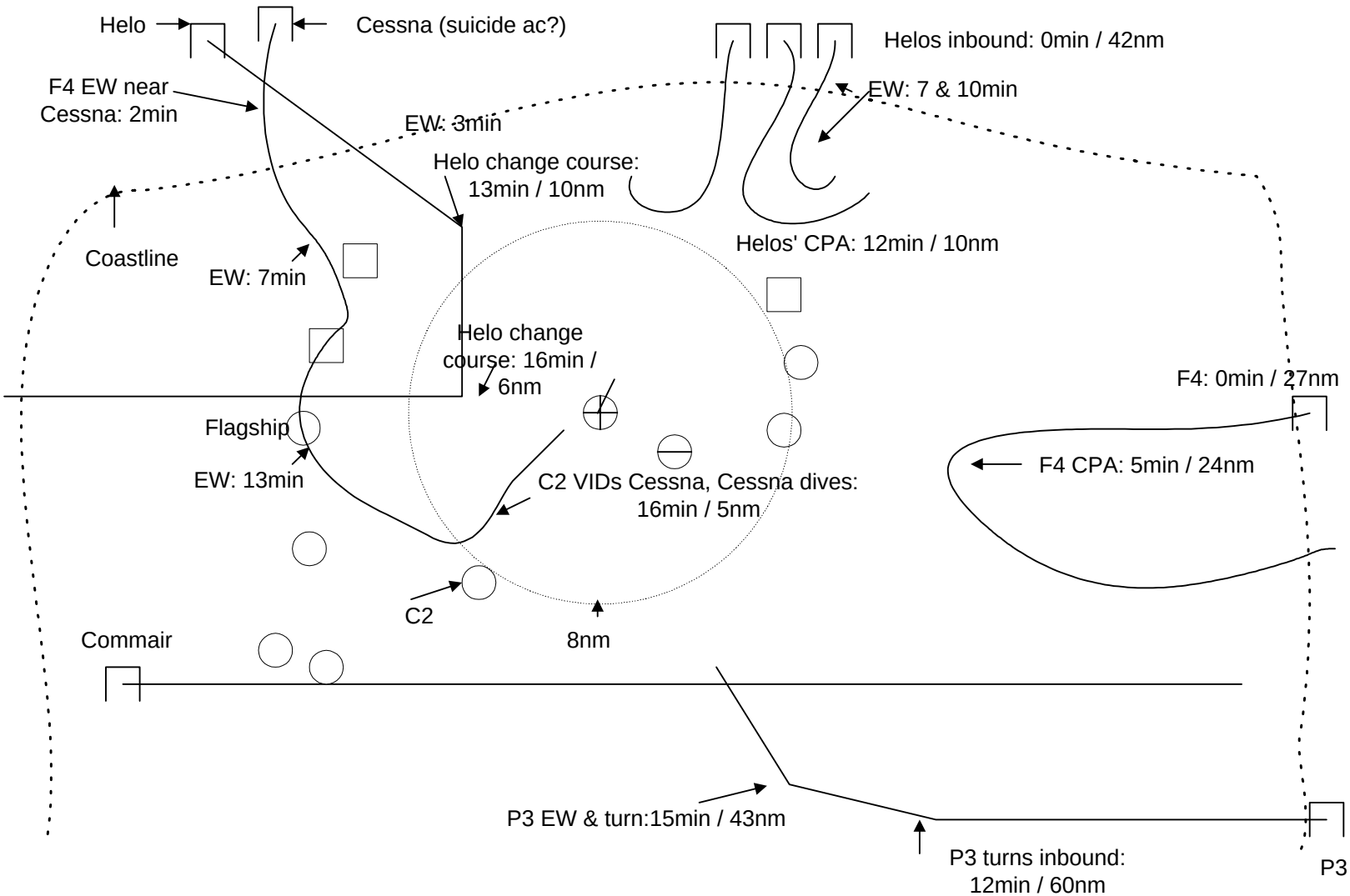
- **Based on decision theoretic Value of Information concept, extended to the value of further attention, or thinking about a problem**
- **Generalized to case where possible outcomes of observations / thinking are unspecified or only vaguely known**
- **Manages thinking process**
  - **Available time = costs of delay in thinking more**
  - **Stakes = costs of potential errors in acting on current recognition**
  - **Uncertainty = match of situation to recognitional patterns, and specific problems (gaps, conflicts, unreliability)**

C T I

Appendix D.
Scenarios and Simulation

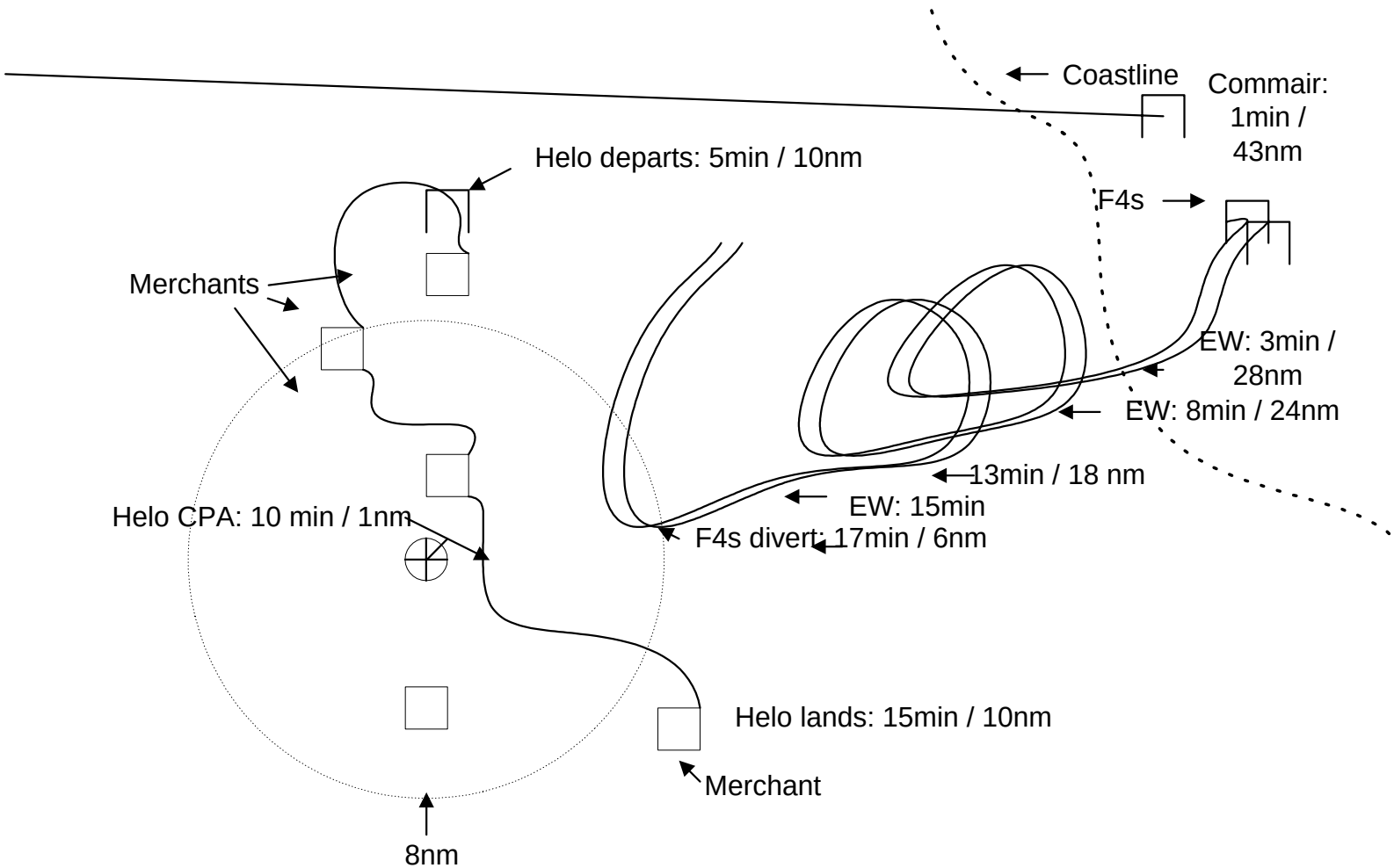
C T I

Scenario Delta



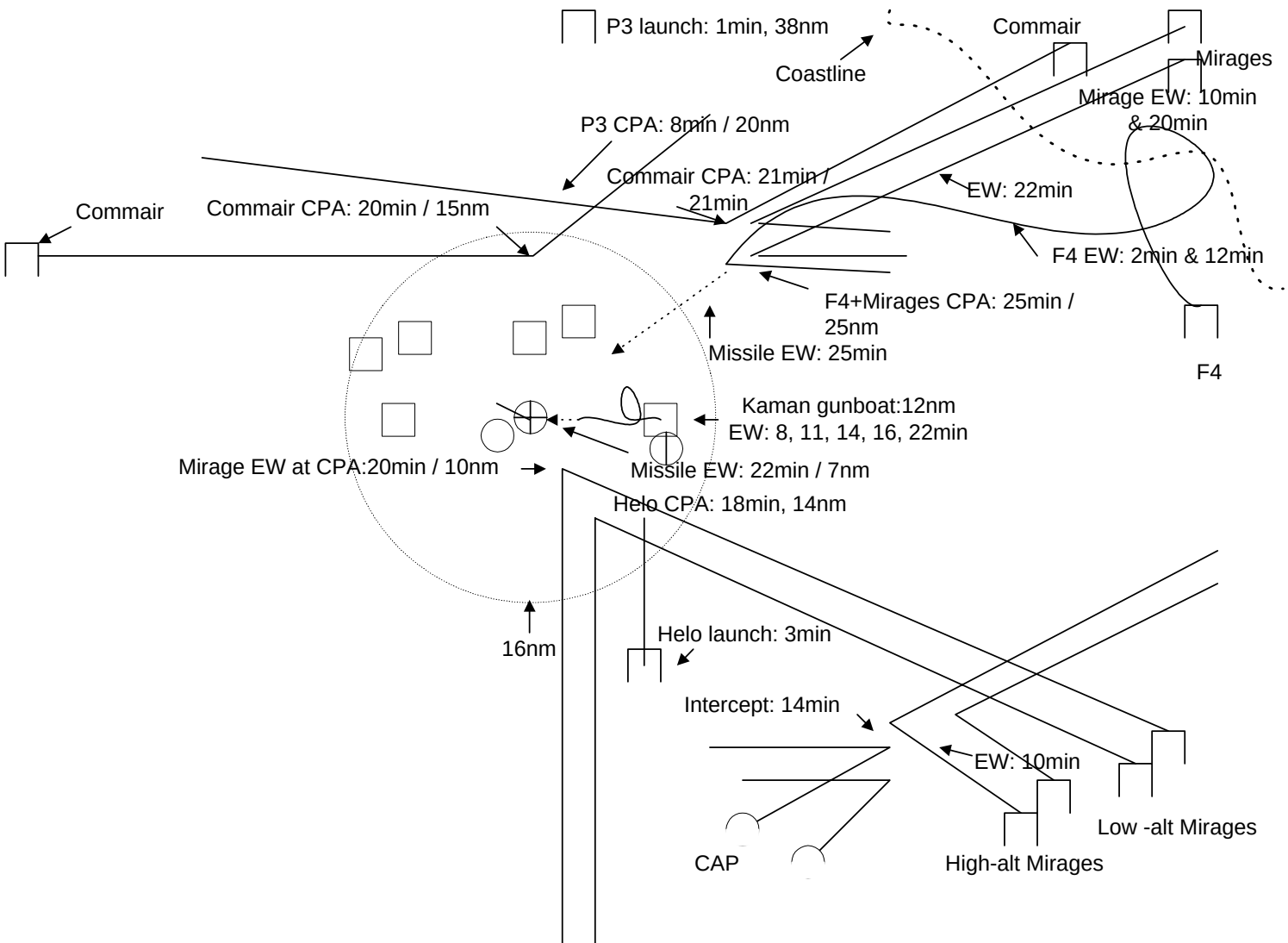
C T I

Scenario Tango



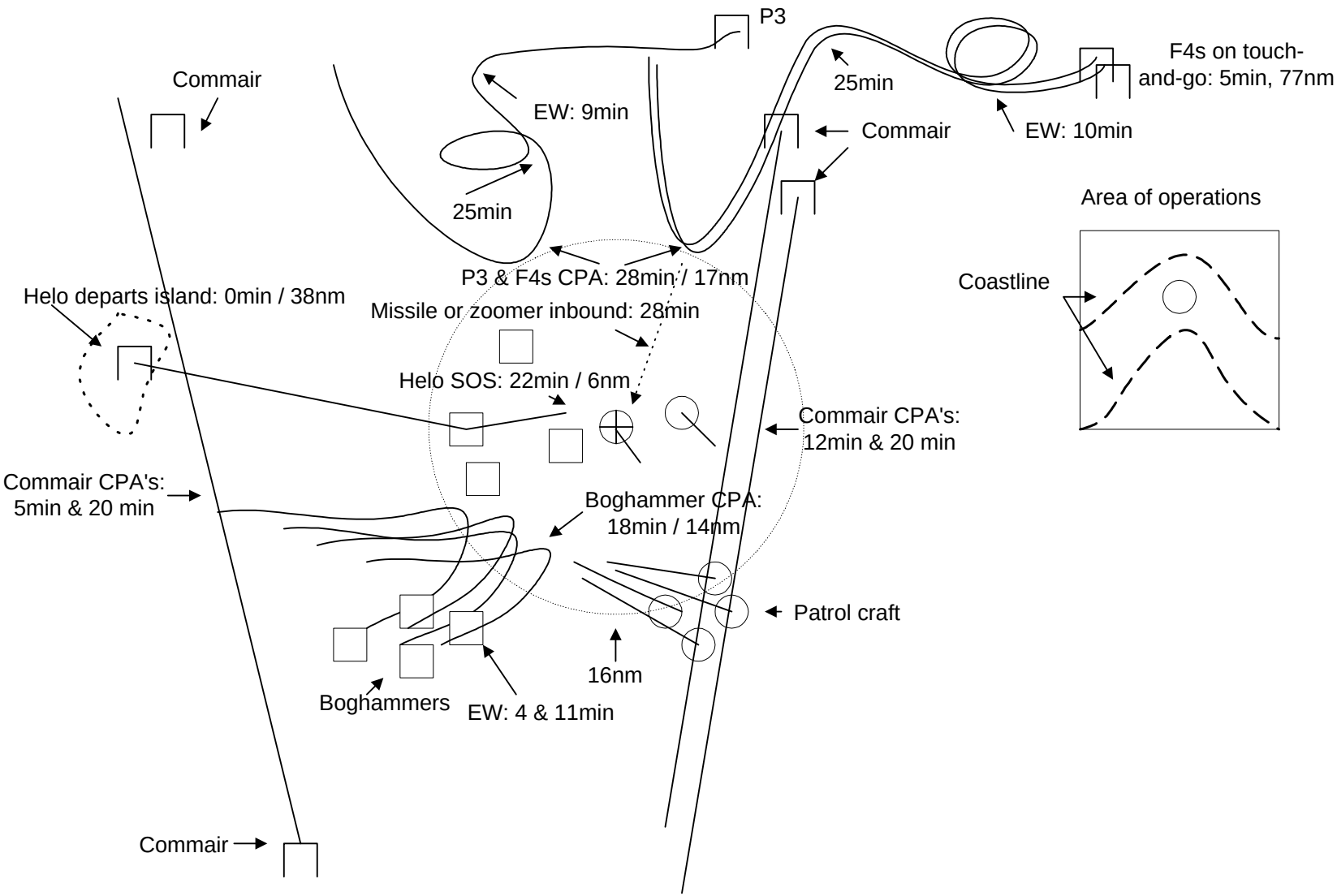
C T I

Scenario Bravo



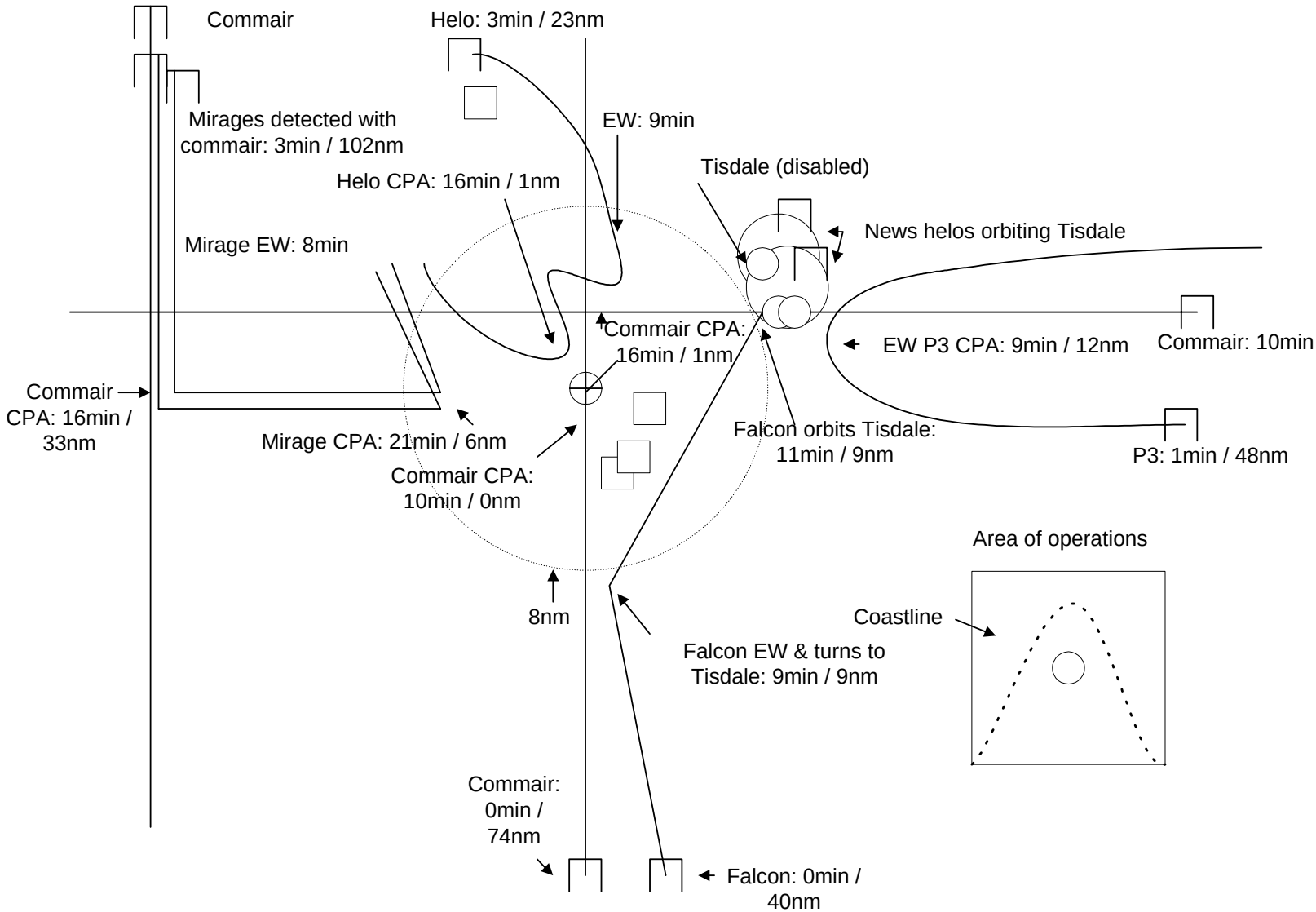
C T I

Scenario Charlie



C T I

Scenario India



Appendix E.
Behavioral Testing

C T I

5. Behavioral Testing

- **Developed efficient measures of critical thinking processes**
- **Developed methods for manipulating metacognitive processing through training**
- **Collected baseline data on human performance for comparison to machine**
- **Tested hypotheses regarding track prioritization by Naval officers**

C T I

Training Metacognitive Skills for Decision Making

Four segments

- **STEP: A Critical Thinking Method**
- **The Hostile Intent template**
- **The Crystal Ball technique**
- **The Quick Test: When to think more**

Each segment includes:

- **Presentation by instructor**
- **Discussion**
- **Exercises with simulated scenarios**
  - » **Continuous real-time probes, discussion, and feedback with instructor**

C T I

Questions at Each Break in Test Scenarios

1 Assess intent of designated tract* and give reasons

2 Give other possible intents and confidence in each

3 Select a disbelieved intent and give reasons it could be true

4 Identify evidence that conflicts with a stipulated intent. Try to explain the evidence.

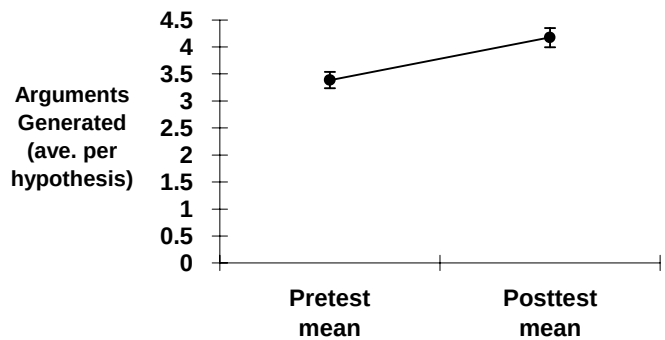
- May or may not be the subject's favored intent in question 1

5 What actions would you take?

* Designated track not known by participants before break

Filling Gaps in Stories

- **Trained participants generated more arguments for a hypothesis**
  - **30% increase (p=.001)**
- **Training increased reasoning based on <u>factors other than track kinematics</u>: i.e., prior situation & motives, opportunities, capability, localization**
- **Training increased arguments based on <u>predicted</u> actions**



C T I

Potential Tests of Metacognitive Learning

Handling conflicting data

- **Present sequences of cues supporting hostile or friendly intent**
  - **Group 1: H1 H2 H3 F1 F2 F3**
  - **Group 2: F1 F2 F3 H1 H2 H3**
- **"Confirmation bias"**
  - **Group 1 interprets F1 as hostile**
  - **Group 2 interprets H1 as friendly**
- **R / M model accounts for this as a form of explanation-based reasoning**
  - **Reinterpret cues in order to construct a single, coherent story (critique reliability of conflicting arguments)**
- **According to R / M, if too many unreliable explanations are required, story is rejected**

C T I

Handling Conflicting Data

Independent variables

- **Order of cues**
- **Number of conflicting cues before engagement is possible**
  - **The more explanations required, the more likely story is to be rejected**
- **Susceptibility of F cues or H cues to a single explanation (e.g., same source, change in tactics or conditions over time)**
  - **If conflicting cues are subject to a single explanation, more are required to cause change of mind**
- **Sequential vs. simultaneous presentation of conflicting cues**
  - **If conflicting cues are sequential, more ability to construct explanations and less ability to evaluate overall reliability of story**
- **Experience in the domain**

C T I

Metacognitively Driven Domain Learning

- **Test learning about cues in a subsequent closely related scenario**
  - **Group 1 more likely to interpret F1 as hostile**
  - **Group 2 more likely to interpret H1 as friendly**
- **Contrast training with novel problems versus training with a consistent problem set**
- **Training with novel problems builds metacognitive skill**